

DOCUMENT RESUME

ED 420 717

TM 028 439

AUTHOR Fouladi, Rachel T.
TITLE Covariance Structure Analysis Techniques under Conditions of Multivariate Normality and Nonnormality - Modified and Bootstrap Based Test Statistics.
PUB DATE 1998-04-00
NOTE 28p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Diego, CA, April 13-17, 1998).
PUB TYPE Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Computer Software; *Maximum Likelihood Statistics; *Monte Carlo Methods; *Multivariate Analysis; Tables (Data)
IDENTIFIERS AMOS Computer Program; Bootstrap Methods; *Covariance Structural Analysis; EQS Computer Program; *Nonnormal Distributions; Resampling Techniques; Type I Errors

ABSTRACT

Covariance structure analytic techniques have become increasingly popular in recent years. During this period, users of statistical software packages have become more and more sophisticated, and more and more researchers are wanting to make sure that they are using the "best" statistic, whether it be for small sample considerations or for issues of nonnormality. At present, none of the confirmatory structure analytic programs include small sample modifications such as the k-factor Bartlett multiplier (1950) or the Swain multiplier (A, Swain, 1975). They do however include modifications to address distributional nonnormality. EQS offers the Satorra-Bentler scaled statistics (A. Satorra and P. Bentler, 1988, 1989); it does not yet offer the Satorra-Bentler adjusted statistic. AMOS, on the other had, offers a bootstrap alternative, however, it does not yet offer either of the Satorra-Bentler modified statistics. This Monte Carlo study addresses whether resampling-based procedures provide improved Type I error control over the modified test statistics such as the k-factor Bartlett modified, Swain modified, Satorra-Bentler scaled, and Satorra-Bentler adjusted test statistics. The study provides evidence on the relative performance of the Beran-Strivastave bootstrap procedure and demonstrates that this procedure does not provide as good control of Type I error under conditions of extremely mild distributional nonnormality as the 0-factor Bartlett or Swain modified maximum likelihood procedure. It does show improved Type I error control over the standard maximum likelihood procedure. (Contains 7 tables, 6 figures, and 42 references.) (Author/SLD)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

Covariance structure analysis techniques under conditions of multivariate normality and nonnormality – Modified and Bootstrap based test statistics

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

Rachel
Fouladi

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)

Rachel T. Fouladi

Department of Educational Psychology
University of Texas at Austin
Austin, Texas 78712-1296

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
- CENTER (ERIC)

- ☒ This document has been reproduced as received from the person or organization originating it.
- ☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Abstract. Covariance structure analytic techniques have become increasingly popular in recent years. During this period, users of statistical software packages have become more and more sophisticated, and more and more researchers are wanting to make sure they are using the "best" statistic, whether it be for small sample considerations or for issues of nonnormality. At present, none of the confirmatory structure analytic programs include small sample modifications such as the k -factor Bartlett multiplier or the Swain multiplier. They do however include modifications to address distributional nonnormality. EQS offers the Satorra-Bentler scaled statistic; it does not yet offer the Satorra-Bentler adjusted statistic. AMOS on the other hand offers a bootstrap alternative, however does not yet offer either of the Satorra-Bentler modified statistics. This Monte Carlo study addresses whether resampling based procedures provide improved Type I error control over the modified test statistics such as the k -factor Bartlett modified, Swain modified, Satorra-Bentler scaled, Satorra-Bentler adjusted test statistics.

Subject descriptors: Covariance structure analysis, bootstrap, resampling, small sample modifications, robustness, multivariate normality, multivariate nonnormality, Type I error.

Introduction

It has long been known that many covariance structure analytic techniques popular with an increasingly large body of researchers in education, the behavioral and social sciences can have extremely poor small sample performance characteristics (e.g., Anderson & Gerbing, 1984; Bearden, Sharma, & Teel, 1982; Boomsma, 1980). It has also long been known that some of these covariance structure analytic techniques are sometimes inappropriate for use under conditions of distributional nonnormality (e.g., Boomsma, 1980; Browne, 1982; Henly, 1993; Hu, Bentler, & Kano, 1992).

Importantly however, as evidenced by regular inquiries on SEMNET, an electronic mail list server dedicated to issues related to structural equation modeling, not all researchers have access to large samples, have data which originate from multivariate normal populations, or have data from populations in which multivariate normality is not a concern (c.f., Anderson & Amemiya, 1985; Browne, 1987; Browne & Shapiro, 1987; Fouladi, 1998, for conditions of asymptotic robustness of normal theory techniques). Fortunately for these researchers, over the years a number of viable alternatives to standard structure analytic techniques have been proposed, some but not all of which have been implemented in statistical software currently on the market.

The alternative covariance structure analysis procedures under consideration in the present paper include multiplicative modifications and the application of computer intensive resampling methods to standard covariance structure analysis techniques. Even though the results of Monte Carlo simulation studies showing that the problems associated with the use of covariance structure analysis techniques under small sample size conditions and/or distributional nonnormality can be greatly alleviated by the use of modified test statistics (Curran, West, & Finch, 1996; Fouladi, 1997a,b, 1998) and resampling based test statistics (Ichikawa & Konishi, 1995) are becoming increasingly well known, until recently the relative performance of the modified procedures was largely undocumented, and to date, the relative performance of modified and resampling based statistics remains unexamined. This paper seeks to fill this void in the literature.

Test Procedures Examined

Consider N independently identically distributed observation vectors obtained from a p -dimensional multivariate population, with non-singular $p \times p$ population covariance matrix Σ . In order to test the hypothesis that a given path, structural, and/or measurement model is an appropriate representation of the pattern of relations among the variables in the population, a researcher can use a variety of structure analysis procedures, most of which are formulated using asymptotic covariance distribution theory.

The goodness of fit test statistics used in covariance structure analysis are generally of the form or some function of $T = cnF$ where c is a constant scaling factor, $n = N - 1$, N is the sample size, and F is the minimum value of $F(S, \Sigma_\theta)$, a scalar valued discrepancy function of the sample covariance matrix S from the population covariance model matrix Σ_θ . The standard parametric covariance structure analysis test statistics implemented in popular software with which one can conduct confirmatory covariance structure analyses, all have $c=1$. Under fairly general conditions, when the null hypothesis is true, the discrepancy function is appropriately chosen and the model is identified, these goodness of fit test statistics are asymptotically chi-square distributed with $g = \frac{1}{2}p(p+1) - q$ degrees of freedom, where p is the number of observed variables and q the number of free parameters in the model (Bentler & Dijkstra, 1985; Browne, 1982, 1984; Satorra and Saris, 1985; Shapiro, 1983, 1986; Steiger, Shapiro, & Browne, 1985).

The poor performance of the standard maximum likelihood statistic and other covariance structure analysis test statistics with $c=1$, under small sample conditions and/or conditions of distributional nonnormality has been widely documented for several decades (e.g., Boomsma, 1980; Henly, 1993). Over the years, in various attempts to address the problems associated with covariance structure analysis under these conditions, a number of alternative procedures have been proposed. The alternative procedures include modified test statistics involving multiplicative scalings of standard covariance structure analysis test statistic suggested by Swain (1975), Satorra and Bentler (1988), and Fouladi (1997c). Other proposed solutions to alleviate problems with standard parametric covariance structure analysis techniques include the application of computer-intensive resampling procedures (Beran & Srivastava, 1985; Bollen & Stine, 1992).

Swain modified test statistic

Swain (1975) proposed the use of a modified maximum likelihood test statistic to obtain improved small sample performance characteristics for the standard maximum likelihood chi-square test statistic. Even though proposed more than two decades ago, described in a widely cited chapter by Browne (1982), resulting in improved small sample performance under conditions of multivariate normality, and easily computed, the Swain modified maximum likelihood test statistic is little known and little used (Fouladi, 1997a,b, 1998). According to Browne, Swain proposed four multipliers which seemed to result in an improvement of the approximation to the chi-square distribution by the standard normal theory covariance structure analysis maximum likelihood goodness of fit statistic under conditions of multivariate normality. Browne provided the one general multiplying factor which appeared slightly preferable to the others, but which only applies to models which are scale invariant. The Swain multiplicative modification has $c = 1 - (t_p - t_y) / (12gn)$, where p is the number of observed variables, $y = (\sqrt{1+8q} - 1) / 2$, $q = p(p+1)/2 - g$, g is the number of degrees of freedom in the model, $n = N - 1$, and $t_x = x(2x^2 + 3x - 1)$.

Fouladi (1996, 1997a,b, 1998) showed that the empirical Type I error rate of the Swain modified maximum likelihood procedure converged more quickly to the nominal level than the standard maximum likelihood implemented in commonly used statistical software packages. Fouladi (1998) also showed that under certain small departures from distributional multivariate normality, the Swain modified maximum likelihood procedure had adequate Type I error control, however, under larger departures from distributional normality exhibited empirical Type I error rates which sometimes (depending on the structure of the data) departed substantially from the nominal level.

k -factor Bartlett multipliers

Bartlett (1950) proposed the usage of $N - p/3 - 2k/3 - 11/6$ instead of $N - 1$ as the multiplicative scaling of the minimum value of the maximum likelihood discrepancy function when testing a k -factor model for p observed variables under reduced sample size conditions, thereby suggesting a scaling of the standard maximum likelihood chi-square statistic by $c = 1 - (2p + 4k + 5) / (6n)$. The modified maximum likelihood test statistic using

this multiplier has been implemented in widely used software with exploratory factor analytic capabilities, though not in software or modules used for confirmatory structure analysis.

Fouladi (1997c) argued that the use of the k -factor Bartlett multiplier could be applied under in confirmatory structure analysis whether or not any latent variables are hypothesized in the model. Fouladi (1997a,b) examined the performance of the 0-factor Bartlett modified maximum likelihood test statistic. Results showed significantly improved Type I error control over the standard maximum likelihood test statistic under conditions of reduced sample size for a variety of models.

As with the Swain modified maximum likelihood procedure, Fouladi (1996, 1997a,b) showed that the empirical Type I error rate of the 0-factor Bartlett modified maximum likelihood procedure converged more quickly to the nominal level than the standard maximum likelihood procedure. Fouladi (1998) also showed that under certain small departures from distributional multivariate normality, 0-factor Bartlett modified had adequate Type I error control, however, under larger departures from distributional normality exhibited empirical Type I error rates which sometimes (depending on the structure of the data) departed substantially from the nominal level.

Satorra & Bentler modified test statistics

Satorra and Bentler (Satorra, 1989; Satorra & Bentler, 1988) presented two approaches to improve the asymptotic statistical behavior of covariance structure analysis statistics. Their methods yield the Satorra-Bentler scaled and adjusted test statistics and are referred to chi-square distributions g and d' degrees of freedom respectively (c.f., Satorra & Bentler, 1988). Though the two approaches were designed primarily to address departures from the reference distribution under conditions of multivariate nonnormality, they have also been shown to be useful under conditions of multivariate normality.

Simulation studies in the nineties have presented clear evidence on the improved performance characteristics of the Satorra-Bentler scaled maximum likelihood test statistic over the standard maximum likelihood test statistic (e.g., Chou & Bentler, 1995; Curran, West, & Finch, 1996; Hu & Bentler, 1995; Hu, Bentler, & Kano, 1992). Not only has the Satorra-Bentler scaled statistic been shown to be useful under conditions of multivariate nonnormality (Chou & Bentler, 1995; Curran, West, & Finch, 1996; Hu & Bentler, 1995; Hu, Bentler, & Kano, 1992), it has also been shown to be useful under conditions of multivariate normality and reduced sample sizes (Curran, West, & Finch, 1996). However, until recently, the performance characteristics of the Satorra-Bentler adjusted maximum likelihood test statistic was largely undocumented. Fouladi (1997a,b, 1998) examined the performance of the Satorra-Bentler adjusted statistic and determined that, under reduced sample size conditions, not only did it show improved performance over the standard maximum likelihood test statistic, it also showed improved performance over the Satorra-Bentler scaled statistic both under conditions of multivariate normality and nonnormality. Though, under large sample sizes the Satorra-Bentler scaled procedure was preferred.

Resampling based test statistics

There are a number of different implementations of resampling techniques (c.f., Yung & Bentler, 1992, 1996). Beran and Srivastava (1985) and Bollen and Stine (1992) considered the use of bootstrap-corrected techniques in covariance structure analysis. The method described by Beran and Srivastava and Bollen and Stine permits the establishment of an empirical reference distribution which can be used to test the null hypothesis that the model is a good fit to the data. The Beran-Srivastava method involves the initial transformation of the data such that the data conform to the null hypothesis. After transformation of the data, a large number of independent bootstrap samples are generated by sampling with replacement, and the test statistic for each bootstrap sample is calculated. The empirical distribution of the test statistic is recorded. This empirical distribution is then used as the reference distribution to which the observed value from the original data set is compared. Because the null hypothesis of model fit is rejected for large observed values, the bootstrap critical value is the $(1-\alpha)$ -percentile of the empirical distribution. Rejection is obtained if the value of the test statistic based on the original sample exceeds the bootstrap critical value.

Though recommended over a decade ago, the performance characteristics of resampling based model testing procedures have not been widely documented in the context of covariance structure analysis. Results from a recent simulation study on the application of bootstrap methods in factor analysis by Ichikawa and Konishi (1995) suggest that resampling based covariance structure analysis techniques perform well under conditions of nonnormality and moderate to large sample sizes, however may not perform as well under small sample conditions.

Ichikawa and Konishi (1995) conducted a Monte Carlo experiment investigating the use of Beran-Srivastava bootstrap methods in maximum likelihood factor analysis under conditions of multivariate normality and

multivariate nonnormality; they found that tests of model fit based on Beran-Srivastava bootstrap methods performed extremely well under conditions of distributional normality and nonnormality. However, given that their study examined the performance of bootstrap based techniques for orthogonal models and data from elliptical distributions, they suggest that the results of their study should be generalized with caution. They also note that since bootstrap samples involve random sampling of data with replacement, some observations are included more than once in a bootstrap sample, and can result in numerically ill conditioned sample covariance matrices, particularly under small sample conditions. As such, the use of bootstrap based covariance structure analysis techniques may not be advisable for extremely small sample sizes.

Empirical results on the relative performance of the alternative procedures

That Monte Carlo simulation results show that the problems associated with the use of covariance structure analysis techniques under small sample size conditions and/or conditions of distributional nonnormality can be greatly alleviated by the use of modified tests statistics and resampling based test statistics is becoming increasingly well known. Less well-known is the relative performance of these alternative techniques, particularly the relative performance of modified and resampling based techniques.

In the only studies to simultaneously examine the performance of the standard maximum likelihood test statistic and the 0-factor Bartlett, Swain and both Satorra-Bentler modified maximum likelihood test statistics, Fouladi (1997a,b, 1998) found that the application of the 0-factor Bartlett, Swain and Satorra-Bentler multipliers to the standard maximum likelihood test statistic yielded significant improvements in Type I error control under conditions of multivariate normality and nonnormality. Fouladi (1997a) found, under conditions of multivariate normality and small sample sizes, the Satorra-Bentler scaled and adjusted test statistic performed better than the standard maximum likelihood test statistic, however did not yield as good performance as a 0-factor Bartlett modified or Swain modified maximum likelihood procedure. Under conditions of multivariate nonnormality, Fouladi (1997b, 1998) found the Satorra-Bentler scaled and adjusted test statistic performed better than the standard, 0-factor Bartlett modified, and Swain maximum likelihood test statistics under models in which the observed variables were more than moderately nonnormally distributed and could not be said to be derived from a population with an orthogonal latent variable structure, otherwise, the 0-factor Bartlett modified and Swain modified maximum likelihood procedures were preferred.

The purpose of this study

To date no study has simultaneously examined the performance of modified covariance structure analysis test statistics and the application of resampling methods to tests of covariance structure analysis recommended by Beran and Srivastava (1985) and Bollen and Stine (1992). As such the relative performance of Beran-Srivastava resampling methods in covariance structure analysis procedures and modified test statistics remains a question. This paper aims to address this unknown using a Monte Carlo simulation project, and in so doing seeks to determine which procedures yield improved Type I error control under conditions of multivariate normality and under conditions of multivariate nonnormality.

Methods

A series of Monte Carlo simulation experiments were conducted in order to compare the error rate control of the standard, 0-factor Bartlett modified, Swain modified, Satorra-Bentler scaled and adjusted test statistics, and the Beran-Srivastava bootstrap covariance structure analysis test procedure under conditions of multivariate normality and nonnormality. The simulation experiment was conducted using a stand-alone FORTRAN computer program implementing the procedures under study. Programming accuracy checks were conducted with Mathematica (Wolfram, 1996), EQS (Bentler, 1995), and the SePath (Steiger, 1995).

For the examination of Type I error control, data sets were generated from 6-dimensional populations with specified moments. The univariate moments of the 6 variables in each multivariate population were homogeneous. In these populations, the means of the variables were all 0, and the variances were all 1. Nine univariate distribution types of varying nonnormality were considered. The distributional characteristics of the variables included levels of skew $\gamma_1=0, 1, 2$, and levels of kurtosis $\gamma_2 = -1, 0, 1, 3, 6$; however, not all possible combinations of skew and kurtosis were examined. The possible combinations of skew and kurtosis are restricted according to the inequality: $\gamma_1^2 < .629576\gamma_2 + .717247$ (Fleishman, 1978). Table 1 details the combinations of univariate skew and kurtosis considered in the present paper.

The underlying correlation structure of the variables was also varied. The correlational structure of the variables was either diagonal (uncorrelated variables) or simplex (correlated variables). Figure 1 details the characteristics of the models in the population using SEPATH path language; note that no latent variables were described in the models.

In the present study, the Vale and Maurelli (1983) method was used to generate data from the specified p -variate populations. This method is a p -variate extension of the Fleishman (1978) method for generating univariate nonnormal data, and can be used to generate data from populations with specified marginal means, variances, skew, and kurtoses and specified correlation structure.

Multivariate data sets were generated at various sample sizes, $N : 2q, 4q, 10q, 20q$, and $50q$, where $q = p(p+1)/2 - g$. Sample covariance matrices were obtained for each of the data sets, and structural hypotheses were tested at two levels of nominal alpha: $\alpha_{\text{Nominal}} = .05$ and $.01$. For each sample covariance matrix, the six structure analysis procedures were conducted, the decisions for the tests were recorded at each of the nominal levels. Resampling based decisions were based on 1000 bootstrap samples.

Experiments under each condition were replicated 5000 times.

Measures of performance

There are different methods of examining the performance of test procedures. When the test statistics have a common reference distribution as is the case with some of the covariance structure analysis test statistics, one can examine the convergence of the distribution of the test statistics in covariance structure analysis to the reference distribution by examining the convergence of the moments of the sampling distribution of the test statistics to the moments of the reference distribution. This method has become popular in recent years, however, in general the only moment that is examined is the mean; little attention has been paid to higher moments. This method is less useful when test statistics have different reference distributions, as is the case in the present paper. An alternative method of examining the performance of test procedures that does not involve the complications of different reference distributions is to examine the convergence of the empirical Type I error rates to the nominal level.

For each condition, the number of rejections obtained for each correlation pattern test are tabulated and transformed into proportion rejected. Rejection rates are based on the number of replications for which convergence obtained. Under each condition, the empirical rejection rate, $\alpha_{\text{Empirical}}$, for each statistic is observed. For each cell, the percent bias ($B\%$) of the observed empirical rejection rate from the expected rejection rate, α_{Nominal} , is obtained where $B\% = 100(\alpha_{\text{Empirical}} - \alpha_{\text{Nominal}}) / \alpha_{\text{Nominal}}$. Empirical performance of the procedures is examined using guidelines set forth by Bradley (1978) and Robey and Barcikowski (1992) for what constitutes acceptable departures from nominal alpha. Null-consistent chi-square goodness of fit values based on a normal approximation to the binomial are also computed. Appropriate summing of the chi-square values can be used to provide tests of the overall control of empirical rejection rates at the nominal level. Multivariate and univariate analysis of variance designs are used to further analyze the percent bias results. In the first set of multivariate and univariate analyses, only the intercept effects in the general linear models are tested, thereby providing tests of the departure of the mean percent bias from 0. In the second set of multivariate and univariate analyses, a four-way factorial design is used to determine the influence of Model type, Distribution type, Sample size, and Nominal alpha on the percent bias.

Results

Table 2 details the empirical rejection rates of the 6 covariance structure analysis procedures examined in this study.

Judgments of overall Type I error control using Bradley-Robey-Barcikowski criteria

Empirical rejection rates are examined using the Bradley, Robey, and Barcikowski (BRB) guidelines for what constitutes acceptable levels of departure of empirical Type I error rates from the nominal level. Bradley (1978) asserted that many researchers are unreasonably generous when defining acceptable departures of empirical alpha from the nominal level. He held that the departure of empirical alpha from the nominal level was "negligible" if empirical alpha was within $\alpha \pm \frac{1}{10}\alpha$ according to a 'fairly stringent criterion', and $\alpha \pm \frac{1}{2}\alpha$ according to the "most liberal criterion that [he] was able to take seriously" which in the remainder of his article he referred to as the 'liberal

criterion'. Robey and Barcikowski (1992) supplement the guidelines provided by Bradley for defining acceptable departures from the nominal level, providing an 'intermediate criterion' of $\alpha \pm \frac{1}{4}\alpha$, and a 'very liberal criterion' of $\alpha \pm \frac{3}{4}\alpha$. This set of guidelines $\{\alpha \pm \frac{1}{10}\alpha, \alpha \pm \frac{1}{4}\alpha, \alpha \pm \frac{1}{2}\alpha, \alpha \pm \frac{3}{4}\alpha\}$ for empirical rates are hereafter referred to as the BRB criteria for empirical alpha.

Empirical rejection rates in Table 2 less than the lower limit of $\alpha \pm \frac{3}{4}\alpha$ are indicated, as are empirical rejection rates that exceed the upper limit of $\alpha \pm \frac{3}{4}\alpha$. Inspection of this table reveals that no procedure consistently provides control of empirical rejection rates within the most liberal of the BRB guidelines for acceptable Type I error control.

Table 3 provides an overview of the proportion of conditions in which the empirical rejection rate is within each of the BRB criteria for empirical alpha, is less than the lower limit of $\alpha \pm \frac{3}{4}\alpha$, and is greater than the upper limit of $\alpha \pm \frac{3}{4}\alpha$. The results in this table indicate that no procedure consistently controls empirical rejection rates within the BRB criteria. The standard maximum likelihood procedure has empirical rejection rates that are within the most liberal criterion for empirical alpha, $\alpha \pm \frac{3}{4}\alpha$, in 39% of the conditions examined. In contrast, the other procedures have empirical rejection rates that are within $\alpha \pm \frac{3}{4}\alpha$ in over 50% of the conditions. Of particular note is that the Satorra-Bentler adjusted maximum likelihood procedure, which shows the best of control of empirical alpha within $\alpha \pm \frac{3}{4}\alpha$, demonstrates control in 83% of the conditions, followed by the Beran-Srivastava bootstrap maximum likelihood procedure which provides control in 66% of the conditions. The Swain and the Satorra-Bentler scaled maximum likelihood procedures provide control within $\alpha \pm \frac{3}{4}\alpha$ in 59% and 52% of the conditions, respectively. The magnitude of the percentages can be used to rank order the procedures from best to worst Type I error control. Similar rankings can be obtained from the percentages within the more stringent BRB criteria for empirical Type I error control. The results obtained from a ranking of the percentages within $\alpha \pm \frac{3}{4}\alpha$ and $\alpha \pm \frac{1}{2}\alpha$ suggest the same overall ordering of the procedures' control of Type I error from best to worst: Satorra-Bentler adjusted, Beran-Srivastava Bootstrap, 0-factor Bartlett modified, Swain modified, Satorra-Bentler scaled, followed by the standard maximum likelihood procedure. A different order is obtained from a ranking of the percentages within $\alpha \pm \frac{1}{4}\alpha$ and $\alpha \pm \frac{1}{10}\alpha$: 0-factor Bartlett modified, Swain modified, Beran-Srivastava bootstrap, Satorra-Bentler adjusted, Satorra-Bentler scaled, followed by the standard maximum likelihood procedure.

Of further note, is that all the procedures with the exception of the Beran-Srivastava bootstrap maximum likelihood procedures had empirical rejection rates in excess of the upper bound of $\alpha \pm \frac{3}{4}\alpha$ a larger proportion of the time than they had rejection rates that were less than the lower bound of $\alpha \pm \frac{3}{4}\alpha$, thereby showing a general tendency of liberal bias over conservative bias for standard maximum likelihood and modified maximum likelihood test procedures. By contrast, the bootstrap procedure showed a general tendency of conservative bias over liberal bias.

Overall chi-square goodness of fit tests on empirical alpha

Null-consistent chi-square goodness of fit values based on a normal approximation to the binomial were computed comparing the empirical rejection rates to nominal alpha. Appropriate summing of the chi-square values was used to provide tests of the overall control of empirical rejection rates at the nominal level. The results showed that none of the procedures provide overall control empirical rejection rates at the nominal level or at or below the nominal level ($p < .001$). The magnitude of the chi-square values can be used to rank order the procedures from best to worst Type I error control. These results suggested the overall the ranking of the procedures' control of Type I error at the nominal level from best to worst is: Beran-Srivastava bootstrap, Satorra-Bentler adjusted, 0-factor Bartlett modified, Swain modified, standard, followed by the Satorra-Bentler scaled maximum likelihood procedure.

Analysis of variance

A multivariate analysis of variance was conducted to test whether the mean percent bias of the 6 procedures was within sampling error of 0. Results showed that overall there was a significant difference between the mean percent bias of the procedures and 0 ($p < .001$). Univariate analyses were conducted for each of the procedures. Results from the univariate analyses showed that, with the exception of the mean percent bias of the Satorra-Bentler

adjusted maximum likelihood procedure ($p=.114$), the mean percent bias of all the other procedures was significantly different from 0 ($p<.001$). Table 4 presents summary statistics on each of the procedures. These mean and standard error of the mean results show that overall the standard, 0-factor Bartlett modified, Swain modified, and Satorra-Bentler scaled maximum likelihood procedures are liberal procedures, and the Beran-Srivastava Bootstrap maximum likelihood procedure is conservative.

A factorial multivariate analysis of variance was conducted to determine the influence of Model type, Distribution type, Sample size, and Nominal alpha on percent bias. With the exception of Model $\times$ Distribution type $\times$ Nominal alpha (Pillai-Bartlett trace, $p=.026$) and Distribution type $\times$ Sample size $\times$ Nominal Alpha (Pillai-Bartlett trace, $p=.152$), all multivariate tests of main effects, two-way and three-way interaction effects yielded $p<.001$; the four way interaction effect was not tested as there was only one summary empirical rejection rate per cell. Univariate factorial analyses were conducted for each test procedure: eta-squared values and p-levels associated with each effect are summarized in Table 5, as are estimated R-squared values ($SS_{\text{Effect}}/SS_{\text{Total}}$). These univariate results show that even though many of the effects in the model are statistically significant, many of the effects account for less than ten percent of the variability in the percent bias. The effects which account for more than ten percent of the variability in the percent bias are: (a) Distribution type and Model $\times$ Distribution type for the standard maximum likelihood procedure, (b) Model type, Distribution type, and Model $\times$ Distribution type for the 0-factor Bartlett and Swain modified maximum likelihood procedures, (c) Sample size and Model $\times$ Sample size for the Satorra-Bentler scaled and adjusted maximum likelihood procedures, (d) Model, Sample size, and Distribution type for the Beran-Srivastava Bootstrap based maximum likelihood procedure.

Type I error control as a function of Model type, Sample size, and Distribution type

Tables 6 and 7 provide summary chi-square goodness of fit values for the control of Type I error rates at the nominal level overall and as a function of model type, sample size, and distribution type. Comparison of the magnitudes of the chi-square values permits a ranking of the procedures in terms of Type I error control across the various conditions. Examination of the overall chi-square values suggests that even though none of the procedures provided Type I error control in general the Beran-Srivastava bootstrap procedure had the best performance. Within the uncorrelated variables model, the 0-factor Bartlett modified procedure outperformed all other procedures across all the distribution types and at $N:q$ of 2, 4, and 10, however at $N:q$ of 20 and 50 the Beran-Srivastava bootstrap procedure had the best Type I error control. Within the correlated variables model, the 0-factor Bartlett modified procedure outperformed the other procedures for the (K,S) equal to (-1,0), (0,0), and (1,0), for the distributions with increased leptokurtosis and/or skew alternative procedures were preferred. The Beran-Srivastava bootstrap procedure outperformed all other procedures at $N:q$ of 2, 10, and 20. At $N:q$ of 4 the Satorra-Bentler adjusted procedure was preferred. At $N:q$ of 50, the Satorra-Bentler scaled procedure was preferred.

Figures 2-3 depict the influence of Model type, Distribution type and Sample size on the empirical rejection rates of the 6 procedures at the .05 nominal level. Figure 2 illustrates how for the uncorrelated variables model, the empirical rejection rates of the procedures vary relatively little as a function of distribution type, and vary mainly as a function of sample size. By contrast, Figure 3 illustrates how for the correlated variables model, the empirical rejection rates of the standard, 0-factor modified, and Swain modified maximum likelihood procedures vary mainly as a function of distribution type, whereas the Satorra-Bentler scaled and adjusted maximum likelihood procedures and the Beran-Srivastava Bootstrap maximum likelihood procedure vary mainly as a function of sample size, and relatively little as a function of distribution type.

Figure 4 illustrates the empirical rejection rates as a function of sample size under conditions of the uncorrelated variables model and nominal alpha equal to .05. The minimal variability in the empirical rejection rates for each procedure at each level of sample size highlights that distribution type is not a major factor under conditions of the uncorrelated variables model. This figure also illustrates the rapid convergence of empirical rejection rates to the nominal level for the 0-factor Bartlett and Swain modified maximum likelihood procedures. At all levels of sample size, the 0-factor Bartlett modified maximum likelihood procedure has the best Type I error control of all the procedures, followed by the Swain modified maximum likelihood procedure. The Beran-Srivastava bootstrap and Satorra-Bentler adjusted procedures showed better Type I error control than the standard maximum likelihood procedure, but not as good as the 0-factor Bartlett modified or the Swain modified procedures.

Figure 5 illustrates the empirical rejection rates as a function of sample size under conditions of the correlated variables model and nominal alpha equal to .05. The differential variability in the empirical rejection rates for some of the procedures at each level of sample size highlights that distribution type is a major factor under conditions of the correlated variables model for the standard, 0-factor Bartlett modified, and Swain modified

maximum likelihood procedures. This figure also illustrates the rapid convergence of empirical rejection rates to the nominal level for the Satorra-Bentler adjusted and Beran-Srivastava bootstrap procedure. At the lowest level of sample sizes, these procedures outperformed all the other procedures including the Satorra-Bentler scaled procedure in Type I error control.

Figure 6 illustrates the empirical rejection rates as a function of distribution type under conditions of the correlated variables model and nominal alpha equal to .05. The differential variability in the empirical rejection rates for some of the procedures at each level of distribution type highlights that distribution type can be a major factor under conditions of the correlated variables model, particularly for the standard, 0-factor, and Swain modified procedures. Importantly, for the first three distribution types, the 0-factor Bartlett and the Swain modified maximum likelihood procedures outperformed all other procedures in Type I error control. Even the standard maximum likelihood procedure performed well under these conditions for moderate to large sample sizes. However for the other distribution types, depending on the level of sample size, the preferred procedures were the Satorra-Bentler scaled and adjusted maximum likelihood procedures or the Beran-Srivastava bootstrap maximum likelihood procedure..

Discussion

Researchers in education, the behavioral and social science are making increasingly frequent use of covariance structure analytic techniques to answer questions of substantive interest. Fortunately, researchers are more and more aware of some of the problems that can be associated with covariance structure analysis under small sample conditions and/or under conditions of nonnormality.

A wide variety of procedures have been proposed to address some of the problems encountered by researchers engaging in research using small data sets or data sets which can not be described as originating from multivariate normal populations. A number of simulation results have evidenced the improved performance of some of these alternative procedures over the standard maximum likelihood test statistic.

Recent simulation results (Fouladi, 1998) evidenced that the covariance structure analysis procedures with the best small sample Type I error control under conditions of extremely mild distributional nonnormality include the 0-factor Bartlett modification or the Swain modification to standard maximum likelihood covariance structure analysis test statistic. The current study provides evidence on the relative performance of the Beran-Srivastava bootstrap procedure and demonstrates that the Beran-Srivastava bootstrap procedure does not provide as good control of Type I error under conditions of extremely mild distributional nonnormality as the 0-factor Bartlett or Swain modified maximum likelihood procedure, however does show improved Type I error control over the standard maximum likelihood procedure.

The procedures with the best Type I error control under more general nonnormal distributional conditions depends largely on the structure of the underlying variables. Fouladi (1998) showed that if the observed variables can be described as originating from populations in which latent variables are orthogonal (as in the case of the uncorrelated variables model), then the 0-factor Bartlett and Swain modified maximum likelihood procedures are preferred; however, if the observed variables can *not* be described as originating from populations in which latent variables are orthogonal (as in the case of the correlated variables model). The alternative structure analytic procedures with some of the best Type I error control under more general nonnormal distributional conditions have been shown to be the Satorra-Bentler adjusted and scaled procedures, with the Satorra-Bentler adjusted procedure providing the best control under conditions of reduced sample size (Fouladi, 1998). The current study provides evidence on the relative performance of the Beran-Srivastava bootstrap procedure, and shows that it does not outperform the 0-factor Bartlett modified or the Swain modified procedures under conditions of mild nonnormality or Satorra-Bentler adjusted procedure under more general conditions of nonnormality. It does however outperform the Satorra-Bentler scaled procedure by showing more rapid convergence of empirical Type I error to the nominal level. Though clearly the Satorra-Bentler scaled procedure outperforms all the procedures under conditions of multivariate nonnormality when sample size is large.

Importantly, however, the choice should not just depend on the performance characteristics of the test procedure, the choice should ultimately be guided by a joint examination of the performance characteristics of the procedures and whether the researcher is in an accept-support or reject-support research situation (Fouladi, 1998; Steiger & Fouladi, 1997; Tanaka, 1987). When researchers using structure analysis are in an accept-support position, that is, researchers are wanting to confirm that the hypothesized model is a good reflection of the population structure, the use of a conservative procedure to fail to reject the null hypothesis that our structure model is a good

reflection of the population structure is inappropriate. Given the Beran-Srivastava procedure is by-and-large the most conservative of the available procedures under conditions of multivariate normality and nonnormality, if the researcher fails to reject the null hypothesis there may be cause for concern, however if the researcher actually manages to reject the null hypothesis with the Beran-Srivastava bootstrap procedure, then the researcher can be satisfied that this has not been done under conditions of elevated Type I error.

By contrast, researchers using structure analysis to refute a model are in a reject-support position, that is, the researchers are wanting to disconfirm that a hypothesized model is a good reflection of the population structure. Under these circumstances, the use of a liberal procedure to obtain a rejection the null hypothesis that the structure model of interest is a good reflection of the population structure is inappropriate. Thus if a researcher rejects the null hypothesis using any of the liberally biased maximum likelihood based procedures there may be cause for concern, however, if the researcher actually fails to reject the null hypothesis then the researcher can be satisfied that this has more than likely been done under conditions of elevated Type I error.

Final note

At present, none of the structure analytic programs commonly used to conduct confirmatory covariance structure analyses include modifications such as the k -factor Bartlett multiplier or the Swain multiplier; this is not of major concern, however, since the modified test statistics resultant from the application of these multipliers are easily obtained. EQS does offer the Satorra-Bentler scaled statistic; however, it does not yet offer the Satorra-Bentler adjusted statistic or the Beran-Srivastava bootstrap procedure. AMOS on the other hand does offer the Beran-Srivastava (Bollen-Stine) bootstrap, however, does not yet offer either of the Satorra-Bentler modified statistics. Even though the other programs do not automatically implement the Beran-Srivastava bootstrap procedure, those programs which enable standard bootstrapping of non-transformed data can be used to obtain proper tests of model fit with some effort.

Author Note

Rachel T. Fouladi (Ph.D., University of British Columbia, 1996) is Assistant Professor in Research Methodology and Data Analysis at the Department of Educational Psychology at the University of Texas at Austin. Correspondence concerning this article should be addressed to the author at: University of Texas at Austin, Dept of Educational Psychology, SZB 504, Austin, TX 78712-1296 U.S.A. Phone 512-471-4155, Fax 512-471-1288, Email rachel.fouladi@mail.utexas.edu, <http://www.edb.utexas.edu/faculty/fouladi/>.

References

- Anderson, J.C., & Gerbing, D.W. (1984). The effect of sampling error on convergence, improper solutions, and goodness of fit indices for maximum likelihood common factor analysis. *Psychometrika*, 49, 155-173.
- Anderson, T.W., & Amemiya, Y. (1985). The asymptotic normal distribution of estimators in factor analysis under general conditions. Technical Report No. 12, prepared under NSF grant DMS 82-19748, Econometric workshop, Stanford University.
- Arbuckle, J. (1996). *AMOS-Analysis of moment structures*. Chicago, IL: Smallwaters Corporation.
- Bearden, W.O., Sharma, S., & Teel, J.R. (1982). Sample size effects on chi-square and other statistics used in evaluating causal models. *Journal of Marketing Research*, 19, 425-530.
- Bentler, P.M. (1995) *EQS structural equations manual*. Encino, CA: Multivariate Software.
- Bentler, P.M., & Dijkstra, T. (1985). Efficient estimation via linearization in structure models. In *Multivariate analysis VI*, ed. P.R. Krishnaiah, pp. 9-42, Amsterdam, North Holland.
- Beran, R., & Srivastava, M.S. (1985). Bootstrap tests and confidence regions for functions of a covariance matrix. *The Annals of Statistics*, 13, 95-115.

- Bollen, K.A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. In *Sociological methodology 1990*, C.C. Clogg, ed., pp. 115-140. Oxford: Basil-Blackwell.
- Boomsma, A. (1983). On the robustness of LISREL (maximum likelihood estimation) against small sample size and nonnormality. Unpublished doctoral dissertation, University of Groningen.
- Bradley, J.V. (1978). Robustness? *British Journal of Mathematical and Statistical Psychology*, 34, 144-152.
- Browne, M.W. (1982). Covariance structures. In *Topics in applied multivariate analysis*, ed. D.M. Hawkins, pp. 72-141. Cambridge: Cambridge University Press.
- Browne, M.W., & Shapiro, A. (1987) Robustness of normal theory methods in analysis of linear latent variate models. *British Journal of Mathematical and Statistical Psychology*, 41, 197.
- Browne, M.W. (1984). Asymptotically distribution free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 37, 62-83.
- Chou, C.P., & Bentler, P.M. (1995). Estimates and tests in structural equation modeling. In *Structural equation modeling: Concepts, issues, and applications*, R. Hoyle, ed., pp. 36-55. Thousand Oaks, CA: Sage.
- Chou, C.P., Bentler, P.M., & Satorra, A. (1991). Scaled test statistics and robust standard errors for nonnormal data in covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical and Statistical Psychology*, 44, 347-357.
- Curran, P.J., West, S.G., & Finch, J.F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods*, 1.
- Fleishman, A.I. (1978). A method for simulating nonnormal distributions. *Psychometrika*, 43, 521-532.
- Fouladi, R.T. (1996). A study of procedures to examine correlation pattern hypotheses under conditions of multivariate normality and nonnormality. Ph.D. dissertation, University of British Columbia, Vancouver, B.C., Canada.
- Fouladi, R.T. (1997a). The use of modified covariance and correlation structure analysis procedures under conditions of multivariate normality: Controlling Type I error. Presented at the International Symposium on Contemporary Multivariate Analysis and Its Application, May 19-22, 1997, Hong Kong.
- Fouladi, R.T. (1997b). Type I error control of some covariance structure analysis techniques under conditions of multivariate non-normality. *Computing Science and Statistics*, 29, 178-186.
- Fouladi, R.T. (1997c). Model Fit in Exploratory and Confirmatory Factor Analysis and Structural Equation Modeling under Conditions of Reduced Sample Size: Test Statistics, Point Estimates, and Confidence Intervals. Manuscript submitted for publication.
- Fouladi, R.T. (1998). Performance of modified test statistics in covariance and correlation structure analysis under conditions of multivariate non-normality. *Structural Equation Modeling*. Manuscript accepted for publication.
- Henly, S.J. (1993). Robustness of some estimators for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, 46, 313-338.
- Hu, L., & Bentler, P.M. (1995). Evaluating model fit. In *Structural equation modeling: Concepts, issues, and applications*, R. Hoyle (Ed.), pp. 36-55. Thousand oaks, CA: Sage.
- Hu, L., Bentler, P.M., & Kano, Y. (1992). Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112, 341-362.
- Ichikawa, M., & Konishi, S. (1995). Application of the bootstrap methods in factor analysis. *Psychometrika*, 1, 77-93.
- Robey, R.R., & Barcikowski, R.S. (1992). Type I error and the number of iterations in Monte Carlo studies of robustness. *British Journal of Mathematical and Statistical Psychology*, 45, 283-288.

- Satorra, A. (1989). Alternative test criteria in covariance structure analysis: a unified approach. *Psychometrika*, 54, 131-151.
- Satorra, A., & Bentler, P.M. (1986). Some robustness properties of goodness of fit statistics in covariance structure analysis. *American Statistical Association 1986 Proceedings of the Business and Economics Sections* (pp. 549-554). Alexandria, VA: American Statistical Association.
- Satorra, A., & Bentler, P.M. (1988). Scaling corrections for chi-square statistics in covariance structure analysis. *American Statistical Association 1988 Proceedings of the Business and Economics Sections* (pp. 308-313). Alexandria, VA: American Statistical Association.
- Satorra, A., & Saris, W.E. (1985). Power of the likelihood ratio test in covariance structure analysis. *Psychometrika*, 50, 83-90.
- Shapiro, A. (1983). Asymptotic distribution theory in the analysis of covariance structures. *South African Statistical Journal*, 17, 33-81.
- Shapiro, A. (1986). Asymptotic theory of over parametrized structure models. *Journal of American Statistical Association*.
- Steiger, J.H. (1995). Structural equation modeling. In *Statistica for Windows* (Computer program manual). Tulsa, OK: Statsoft, Inc.
- Steiger, J.H., & Fouladi, R.T. (1997). Noncentrality interval estimation and the evaluation of statistical models. In L. Harlow, S. Mulaik, & J.H. Steiger (Eds.) *What if there were no significance tests*. New Jersey: Lawrence Erlbaum.
- Steiger, J.H., Shapiro, A., & Browne, M.W. (1985). On the multivariate asymptotic distribution of sequential chi-square statistics. *Psychometrika*, 50, 253-264.
- Swain, A.J. (1975). Analysis of parametric structures for variance matrices. Ph.D. dissertation, University of Adelaide.
- Tanaka, J.S. (1987). How big is big enough? Sample size and goodness of fit in structural equation models with latent variables. *Child Development*, 58, 134-156.
- Vale, C.D., & Maurelli, V.A. (1983). Simulating multivariate nonnormal distributions. *Psychometrika*, 48, 465-571.
- Wolfram Research, Inc. (1996). *Mathematica: Win Professional. Version 3.0.1*. Champaign, IL: Wolfram Research, Inc.
- Yung, Y. -F., & Bentler, P. M. (1994). Bootstrap-corrected ADF test statistics in covariance structure analysis. *British Journal of Mathematical and Statistical Psychology*, 47, 63-84.
- Yung, Y. -F., & Bentler, P. M. (1996). Bootstrapping techniques in analysis of mean and covariance structures. In G. A. Marcoulides & R. E. Schumacker (Eds.), *Advanced structural equation modeling techniques* (pp. 195-226). Hillsdale, NJ: LEA.

Tables and Figures

Table 1: Possible combinations of univariate kurtosis and skew

Kurtosis	Skew		
	0	1	2
-1	Distribution 1		
0	Distribution 2		
1	Distribution 3	Distribution 4	
3	Distribution 5	Distribution 6	
6	Distribution 7	Distribution 8	Distribution 9

Table 2: Empirical rejection rates as a function of Model type (uncorrelated variables= 1, correlated variables=2), sample size (N), Distribution type as specified by the marginal kurtoses (Ku) and skew (Sk), and Nominal alpha.

				Alpha=.05					Alpha=.01						
Model	N	Ku	Sk	ML	Bart	Swain	Scl	Adj	Boot	ML	Bart	Swain	Scl	Adj	Boot
1	2q	-1	0	.2514+	.0730	.0838	.5426+	.2118+	.0000-	.1008+	.0186+	.0246+	.3208+	.0420+	.0000-
		0	0	.2624+	.0708	.0856	.5294+	.2204+	.0000-	.1088+	.0150	.0206+	.3308+	.0420+	.0000-
		1	0	.2556+	.0708	.0890+	.5382+	.2064+	.0000-	.1056+	.0182+	.0264+	.3296+	.0352+	.0000-
		1	1	.2510+	.0672	.0814	.5410+	.1976+	.0000-	.1024+	.0140	.0208+	.3214+	.0350+	.0000-
		3	0	.2562+	.0730	.0860	.5448+	.1916+	.0000-	.1042+	.0192+	.0244+	.3218+	.0322+	.0000-
		3	1	.2520+	.0686	.0840	.5234+	.1818+	.0000-	.1032+	.0176+	.0240+	.3152+	.0308+	.0000-
		6	0	.2644+	.0784	.0936+	.5580+	.1782+	.0000-	.1132+	.0182+	.0234+	.3236+	.0288+	.0000-
		6	1	.2564+	.0740	.0896+	.5378+	.1620+	.0000-	.1090+	.0190+	.0232+	.3110+	.0238+	.0000-
	4q	6	2	.2652+	.0870	.1016+	.5434+	.1744+	.0000-	.1216+	.0242+	.0308+	.3182+	.0284+	.0000-
		-1	0	.1106+	.0490	.0558	.2094+	.0716	.0064-	.0272+	.0096	.0100	.0754+	.0062	.0000-
		0	0	.1074+	.0530	.0570	.1978+	.0666	.0102-	.0332+	.0120	.0128	.0704+	.0086	.0012-
		1	0	.1094+	.0558	.0620	.1998+	.0612	.0074-	.0322+	.0100	.0120	.0712+	.0070	.0000-
		1	1	.1054+	.0500	.0540	.2064+	.0626	.0074-	.0284+	.0102	.0114	.0700+	.0068	.0002-
		3	0	.1082+	.0516	.0574	.1980+	.0502	.0078-	.0296+	.0124	.0134	.0674+	.0036	.0002-
		3	1	.1202+	.0590	.0648	.2084+	.0572	.0078-	.0340+	.0128	.0142	.0720+	.0052	.0002-
		6	0	.1224+	.0610	.0650	.1972+	.0434	.0082-	.0376+	.0158	.0172	.0660+	.0024-	.0000-
10q	6	1	.1196+	.0580	.0632	.1944+	.0444	.0092-	.0376+	.0164	.0184+	.0638+	.0038	.0002-	
	6	2	.1212+	.0664	.0704	.1914+	.0392	.0072-	.0430+	.0200+	.0222+	.0588+	.0040	.0002-	
	-1	0	.0702	.0500	.0514	.0952+	.0468	.0388	.0166	.0112	.0124	.0248+	.0072	.0068	
	0	0	.0628	.0450	.0458	.0928+	.0434	.0374	.0142	.0084	.0090	.0226+	.0058	.0046	
	1	0	.0696	.0456	.0482	.0914+	.0372	.0346	.0134	.0088	.0094	.0194+	.0036	.0046	
	1	1	.0662	.0502	.0504	.0880+	.0420	.0380	.0170	.0114	.0122	.0242+	.0058	.0068	
	3	0	.0720	.0528	.0554	.0844	.0362	.0390	.0182+	.0116	.0126	.0218+	.0030	.0046	
	3	1	.0758	.0540	.0554	.0898+	.0360	.0382	.0156	.0108	.0118	.0224+	.0030	.0042	
20q	6	0	.0772	.0598	.0608	.0884+	.0308	.0398	.0218+	.0152	.0158	.0260+	.0030	.0070	
	6	1	.0694	.0548	.0560	.0820	.0250	.0282	.0202+	.0146	.0150	.0170	.0030	.0032	
	6	2	.0772	.0606	.0618	.0844	.0266	.0310	.0236+	.0166	.0166	.0190+	.0024-	.0038	
	-1	0	.0568	.0492	.0504	.0674	.0460	.0464	.0130	.0106	.0110	.0160	.0086	.0100	
	0	0	.0576	.0498	.0504	.0736	.0486	.0488	.0140	.0116	.0118	.0154	.0078	.0088	
	1	0	.0608	.0518	.0534	.0694	.0448	.0504	.0144	.0118	.0120	.0154	.0066	.0102	
	1	1	.0532	.0454	.0466	.0614	.0358	.0412	.0112	.0092	.0094	.0124	.0042	.0074	
	3	0	.0666	.0568	.0576	.0700	.0360	.0486	.0156	.0124	.0126	.0156	.0044	.0092	
50q	3	1	.0568	.0488	.0498	.0630	.0328	.0440	.0086	.0078	.0080	.0110	.0036	.0072	
	6	0	.0590	.0508	.0514	.0648	.0308	.0466	.0156	.0130	.0134	.0130	.0028	.0060	
	6	1	.0706	.0608	.0618	.0714	.0322	.0482	.0164	.0136	.0138	.0142	.0032	.0084	
	6	2	.0662	.0586	.0596	.0680	.0276	.0416	.0186+	.0156	.0158	.0124	.0032	.0058	
	-1	0	.0584	.0564	.0568	.0622	.0536	.0490	.0118	.0108	.0108	.0134	.0090	.0136	
	0	0	.0518	.0468	.0472	.0556	.0456	.0436	.0114	.0106	.0106	.0114	.0074	.0100	
	1	0	.0546	.0520	.0528	.0574	.0480	.0480	.0124	.0114	.0114	.0122	.0074	.0124	
	1	1	.0540	.0500	.0502	.0560	.0446	.0508	.0106	.0100	.0100	.0110	.0070	.0100	
	3	0	.0564	.0502	.0508	.0608	.0408	.0502	.0104	.0098	.0098	.0098	.0054	.0102	
	3	1	.0526	.0500	.0500	.0556	.0406	.0482	.0112	.0100	.0100	.0110	.0052	.0106	
	6	0	.0574	.0540	.0542	.0534	.0354	.0490	.0144	.0130	.0130	.0100	.0050	.0114	
	6	1	.0582	.0540	.0544	.0572	.0384	.0482	.0154	.0146	.0148	.0124	.0054	.0126	
	6	2	.0578	.0556	.0558	.0562	.0318	.0464	.0126	.0112	.0116	.0106	.0036	.0104	

continued

Table 3: Proportion of empirical rejection rates within each BRB criteria for empirical alpha.

	ML	Bart	Swain	Scl	Adj	Boot
$\alpha \pm \frac{1}{10}\alpha$	.056	.256	.206	.100	.150	.194
$\alpha \pm \frac{1}{4}\alpha$	.189	.472	.433	.300	.328	.411
$\alpha \pm \frac{1}{2}\alpha$	.322	.578	.528	.439	.661	.606
$\alpha \pm \frac{3}{4}\alpha$	.394	.628	.589	.522	.833	.661
less than lower limit of $\alpha \pm \frac{3}{4}\alpha$	.000	.000	.000	.000	.011	.289
greater than lower limit of $\alpha \pm \frac{3}{4}\alpha$	.606	.372	.411	.478	.156	.050

Table 4: Summary statistics for the percent bias (B_3) results.

	ML	Bart	Swain	Scl	Adj	Boot
Minimum	-14	-22	-20	-14	-76	-100
Maximum	2848	2798	2798	3208	341	126
Median	142	34	40	68	-18	-17
Mean	350	212	219	343	11	-25
Standard error of the mean	39	33	33	52	7	4
Standard deviation	520	449	448	692	94	56

Table 5 Factorial Analysis of Variance results on percent bias (B_3)

Effect	df _{Effect}	df _{Error}	Eta - squared						R - squared					
			ML	Bart	Swain	Scl	Adj	Boot	ML	Bart	Swain	Scl	Adj	Boot
Model	1	32	.988 a	.980 a	.980 a	.998 a	.965 a	.979 a	.060	.144	.134	.057	.025	.143
N	4	32	.986 a	.586 a	.473 a	1.000 a	.999 a	.995 a	.052	.004	.003	.403	.696	.572
Distribution	8	32	.996 a	.988 a	.988 a	.422 c	.982 a	.532 a	.188	.239	.238	.000	.051	.143
Alpha	1	32	.987 a	.921 a	.928 a	.997 a	.969 a	.161 c	.056	.035	.036	.036	.029	.001
Model $\times$ N	4	32	.979 a	.779 a	.826 a	.999 a	.994 a	.886 a	.033	.010	.013	.138	.161	.024
Model $\times$ Distribution	8	32	.996 a	.986 a	.987 a	.684 a	.216	.692 a	.164	.214	.213	.000	.000	.007
N $\times$ Distribution	32	32	.770 a	.745 a	.752 b	.775 a	.902 a	.658 c	.002	.009	.009	.000	.008	.006
Model $\times$ Alpha	1	32	.952 a	.905 a	.905 a	.993 a	.509 a	.644 a	.014	.028	.027	.017	.001	.006
N $\times$ Alpha	4	32	.919 a	.349 b	.289 c	.999 a	.788 a	.922 a	.008	.002	.001	.100	.003	.037
Distribution $\times$ Alpha	8	32	.987 a	.955 a	.955 a	.357	.616 a	.228	.053	.060	.060	.001	.001	.001
Model $\times$ N $\times$ Distribution	32	32	.813 a	.764 a	.766 a	.783 a	.867 a	.732 a	.003	.010	.009	.000	.006	.008
Model $\times$ N $\times$ Alpha	4	32	.894 a	.447 a	.498 a	.998 a	.341 b	.810 a	.006	.002	.003	.048	.000	.013
Model $\times$ Distribution $\times$ Alpha	8	32	.984 a	.949 a	.951 a	.255	.358	.232	.045	.055	.054	.000	.001	.001
N $\times$ Distribution $\times$ Alpha	32	32	.429	.485	.496	.478	.696 c	.383	.001	.003	.003	.000	.002	.002
Model $\times$ N $\times$ Distribution $\times$ Alpha														

Note: a= $p < .001$, b= $p < .01$, c= $p < .05$

Table 2 - continued: Empirical rejection rates as a function of Model type (uncorrelated variables= 1, correlated variables=2), sample size (N), Distribution type as specified by the marginal kurtoses (Ku) and skew (Sk), and Nominal alpha.

Model	N	Ku	Sk	Alpha=.05						Alpha=.01					
				ML	Bart	Swain	Scl	Adj	Boot	ML	Bart	Swain	Scl	Adj	Boot
1	2q	-1	0	.1140+	.0598	.0612	.2061+	.1086+	.0120-	.0364+	.0140	.0144	.0814+	.0224+	.0012-
		0	0	.1191+	.0594	.0606	.2121+	.1099+	.0128-	.0334+	.0120	.0124	.0841+	.0220+	.0004-
		1	0	.1178+	.0618	.0620	.2056+	.1014+	.0112-	.0342+	.0154	.0156	.0768+	.0188+	.0000-
		1	1	.2263+	.1320+	.1330+	.2155+	.0911+	.0108-	.0829+	.0373+	.0381+	.0801+	.0146	.0006-
		3	0	.1441+	.0817	.0825	.2012+	.0903+	.0116-	.0496+	.0238+	.0244+	.0741+	.0164	.0002-
		3	1	.2132+	.1303+	.1307+	.2294+	.0883+	.0120-	.0795+	.0338+	.0342+	.0803+	.0144	.0006-
		6	0	.2330+	.1489+	.1514+	.2563+	.0955+	.0124-	.0981+	.0491+	.0499+	.0959+	.0160	.0004-
		6	1	.2474+	.1590+	.1598+	.2440+	.0858	.0120-	.1117+	.0596+	.0608+	.0866+	.0132	.0014
		6	2	.4871+	.3706+	.3730+	.2797+	.0853	.0140	.2881+	.1855+	.1877+	.0998+	.0095	.0004
	4q	-1	0	.0800	.0558	.0558	.1136+	.0702	.0372	.0200+	.0120	.0122+	.0328+	.0114	.0060
		0	0	.0704	.0482	.0486	.1040+	.0580	.0302	.0144	.0080	.0082	.0288+	.0088	.0046
		1	0	.0754	.0548	.0552	.1064+	.0562	.0316	.0186+	.0118	.0122	.0264+	.0068	.0050
		1	1	.1840+	.1406+	.1412+	.1174+	.0524	.0362	.0616+	.0412+	.0414+	.0306+	.0062	.0060
		3	0	.1286+	.0946+	.0950+	.1146+	.0542	.0378	.0410+	.0302+	.0306+	.0314+	.0056	.0070
		3	1	.1770+	.1390+	.1404+	.1194+	.0488	.0448	.0698+	.0498+	.0500+	.0314+	.0074	.0084
		6	0	.2246+	.1840+	.1848+	.1234+	.0458	.0470	.1014+	.0716+	.0720+	.0342+	.0044	.0078
		6	1	.2428+	.2022+	.2028+	.1222+	.0414	.0434	.1096+	.0854+	.0858+	.0302+	.0036	.0078
		6	2	.4769+	.4251+	.4263+	.1332+	.0426	.0474	.2815+	.2314+	.2326+	.0362+	.0030	.0056
	10q	-1	0	.0578	.0496	.0496	.0690	.0536	.0508	.0128	.0108	.0108	.0162	.0090	.0102
		0	0	.0582	.0518	.0518	.0696	.0554	.0508	.0134	.0110	.0114	.0160	.0078	.0110
		1	0	.0682	.0602	.0604	.0684	.0498	.0548	.0154	.0124	.0124	.0162	.0088	.0112
		1	1	.1616+	.1438+	.1440+	.0736	.0494	.0612	.0536+	.0462+	.0462+	.0168	.0062	.0116
		3	0	.1192+	.1076+	.1080+	.0692	.0408	.0566	.0354+	.0304+	.0304+	.0144	.0054	.0104
		3	1	.1760+	.1602+	.1608+	.0712	.0394	.0574	.0622+	.0536+	.0536+	.0136	.0042	.0120
		6	0	.2412+	.2228+	.2232+	.0716	.0302	.0670	.1084+	.0974+	.0978+	.0162	.0034	.0146
		6	1	.2680+	.2492+	.2492+	.0678	.0260	.0664	.1244+	.1132+	.1132+	.0118	.0028	.0120
		6	2	.4704+	.4504+	.4514+	.0764	.0304	.0662	.2692+	.2484+	.2490+	.0170	.0034	.0130
20q	-1	0	.0574	.0542	.0542	.0624	.0556	.0530	.0124	.0110	.0110	.0124	.0102	.0118	
	0	0	.0552	.0510	.0514	.0608	.0530	.0530	.0120	.0102	.0102	.0126	.0100	.0126	
	1	0	.0636	.0594	.0596	.0590	.0490	.0516	.0142	.0132	.0132	.0132	.0088	.0120	
	1	1	.1466+	.1398+	.1400+	.0596	.0470	.0540	.0492+	.0466+	.0466+	.0124	.0066	.0112	
	3	0	.1212+	.1144+	.1148+	.0558	.0342	.0510	.0334+	.0314+	.0316+	.0134	.0054	.0120	
	3	1	.1836+	.1744+	.1744+	.0578	.0354	.0556	.0656+	.0624+	.0624+	.0116	.0052	.0106	
	6	0	.2570+	.2478+	.2478+	.0628	.0314	.0678	.1206+	.1148+	.1148+	.0106	.0034	.0174	
	6	1	.2892+	.2784+	.2786+	.0568	.0272	.0600	.1384+	.1324+	.1324+	.0106	.0030	.0118	
	6	2	.4772+	.4662+	.4664+	.0608	.0326	.0604	.2766+	.2682+	.2682+	.0104	.0030	.0138	
50q	-1	0	.0494	.0486	.0486	.0530	.0520	.0580	.0102	.0096	.0096	.0104	.0088	.0178+	
	0	0	.0548	.0534	.0534	.0554	.0528	.0588	.0120	.0118	.0118	.0124	.0108	.0218+	
	1	0	.0606	.0602	.0602	.0542	.0506	.0582	.0144	.0140	.0140	.0116	.0084	.0200+	
	1	1	.1440+	.1418+	.1418+	.0528+	.0474	.0626	.0470+	.0458+	.0458+	.0088+	.0074	.0216+	
	3	0	.1228+	.1192+	.1192+	.0492+	.0400	.0586	.0344+	.0332+	.0332+	.0102+	.0072	.0180+	
	3	1	.1814+	.1778+	.1778+	.0562+	.0432	.0652	.0736+	.0712+	.0712+	.0100+	.0068	.0226+	
	6	0	.2750+	.2724+	.2726+	.0512+	.0324	.0640	.1254+	.1216+	.1216+	.0086+	.0032	.0216+	
	6	1	.3280+	.3242+	.3242+	.0538+	.0338	.0726	.1596+	.1572+	.1572+	.0112+	.0054	.0224+	
	6	2	.4950+	.4904+	.4906+	.0488+	.0374	.0632	.2948+	.2898+	.2898+	.0094+	.0048	.0176+	

Table 6: Summary chi-square goodness of fit tests overall and as function of Model type (uncorrelated variables=1, correlated variables=2) and Sample size (N).

Model	N	df	ML	Bart	Swain	Scl	Adj	Boot
1	2q	18	84199.00 a	900.29 a	2381.16 a	668269.13 a	21882.61 a	2822.97 a
	4q	18	6543.99 a	158.35 a	298.90 a	37004.46 a	250.75 a	2108.93 a
	10q	18	772.22 a	83.24 a	101.99 a	2099.38 a	403.18 a	315.24 a
	20q	18	252.84 a	63.51 a	73.01 a	392.31 a	338.38 a	49.39 a
	50q	18	70.39 a	29.96 c	33.02 c	68.21 a	174.77 a	21.20
	Subtotal	90	91838.44 a	1235.35 a	2888.08 a	707833.49 a	23049.70 a	5317.73 a
2	2q	18	91209.86 a	32974.99 a	33774.34 a	55968.64 a	2259.40 a	1766.70 a
	4q	18	80941.55 a	52100.25 a	52627.13 a	6439.12 a	160.38 a	202.98 a
	10q	18	79107.29 a	66537.91 a	66840.63 a	557.13 a	276.38 a	134.37 a
	20q	18	86195.56 a	79853.36 a	79894.58 a	113.01 a	277.43 a	109.10 a
	50q	18	99669.53 a	96253.08 a	96280.99 a	19.57	156.92 a	670.33 a
	Subtotal	90	437123.78 a	327719.58 a	329417.67 a	63097.46 a	3130.50 a	2883.48 a
Grand Total		180	528962.22 a	328954.93 a	332305.76 a	770930.95 a	26180.20 a	8201.21 a

Note: a=p<.001, b=p<.01, c=p<.05

Table 7: Summary chi-square goodness of fit tests overall and as function of Model type (uncorrelated variables=1, correlated variables=2) and Distribution type.

Model	Ku	Sk	df	ML	Bart	Swain	Scl	Adj	Boot
1	-1	0	10	9053.05 a	98.84 a	240.28 a	79560.77 a	3338.87 a	590.66 a
	0	0	10	10339.14 a	67.62 a	204.29 a	80660.73 a	3625.20 a	556.14 a
	1	0	10	9759.24 a	89.20 a	318.70 a	81217.64 a	2963.77 a	598.25 a
	1	1	10	9114.84 a	42.79 a	169.23 a	79019.42 a	2692.25 a	585.16 a
	3	0	10	9641.98 a	111.53 a	268.82 a	79106.73 a	2481.05 a	577.64 a
	3	1	10	9585.00 a	82.68 a	259.37 a	75483.25 a	2182.29 a	589.54 a
	6	0	10	11342.18 a	183.20 a	382.37 a	81011.80 a	2105.89 a	573.98 a
	6	1	10	10508.27 a	173.27 a	359.09 a	74660.06 a	1611.41 a	616.16 a
	6	2	10	12494.74 a	386.24 a	685.93 a	77113.08 a	2048.97 a	630.19 a
	Subtotal		90	91838.44 a	1235.35 a	2888.08 a	707833.49 a	23049.70 a	5317.73 a
2	-1	0	10	947.76 a	26.82 b	32.05 a	5907.73 a	489.53 a	256.57 a
	0	0	10	855.44 a	17.61	21.10	6101.08 a	465.61 a	331.62 a
	1	0	10	984.74 a	80.84 a	84.51 a	5346.15 a	329.33 a	320.26 a
	1	1	10	14843.06 a	7086.25 a	7151.50 a	6154.37 a	214.99 a	336.08 a
	3	0	10	5359.99 a	2620.69 a	2661.31 a	5207.66 a	274.14 a	270.83 a
	3	1	10	19093.44 a	11464.37 a	11529.32 a	6686.29 a	240.65 a	316.53 a
	6	0	10	46354.80 a	33769.48 a	33961.01 a	9156.62 a	431.56 a	389.46 a
	6	1	10	63648.38 a	49271.54 a	49410.70 a	7724.51 a	373.39 a	370.30 a
	6	2	10	285036.17 a	223381.99 a	224566.17 a	10813.06 a	311.30 a	291.82 a
	Subtotal		90	437123.78 a	327719.58 a	329417.67 a	63097.46 a	3130.50 a	2883.48 a
Grand Total			180	528962.22 a	328954.93 a	332305.76 a	770930.95 a	26180.20 a	8201.21 a

Note: a=p<.001, b=p<.01, c=p<.05

Figure 1: Population covariance matrices and models.

Uncorrelated variables population and model.

Population covariance matrix.

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Model tested, $q=6$.

[VAR1]-{0.0}-[VAR2]
 -(0.0)-[VAR3]
 -(0.0)-[VAR4]
 -(0.0)-[VAR5]
 -(0.0)-[VAR6]

[VAR2]- {0.0}-[VAR3]
 -(0.0)-[VAR4]
 -(0.0)-[VAR5]
 -(0.0)-[VAR6]

[VAR3]- {0.0}-[VAR4]
 -(0.0)-[VAR5]
 -(0.0)-[VAR6]

[VAR4]- {0.0}-[VAR5]
 -(0.0)-[VAR6]

[VAR5]- {0.0}-[VAR6]

[VAR1]- 1 -[VAR1]
 [VAR2]- 2 -[VAR2]
 [VAR3]- 3 -[VAR3]
 [VAR4]- 4 -[VAR4]
 [VAR5]- 5 -[VAR5]
 [VAR6]- 6 -[VAR6]

Correlated variables population and model.

Population covariance matrix.

$$\begin{bmatrix} 1 & .7 & .6 & .5 & .4 & .3 \\ .7 & 1 & .7 & .6 & .5 & .4 \\ .6 & .7 & 1 & .7 & .6 & .5 \\ .5 & .6 & .7 & 1 & .7 & .6 \\ .4 & .5 & .6 & .7 & 1 & .7 \\ .3 & .4 & .5 & .6 & .7 & 1 \end{bmatrix}$$

Model tested, $q=11$.

[VAR1]-1-[VAR2]
 -2-[VAR3]
 -3-[VAR4]
 -4-[VAR5]
 -5-[VAR6]

[VAR2]-1-[VAR3]
 -2-[VAR4]
 -3-[VAR5]
 -4-[VAR6]

[VAR3]-1-[VAR4]
 -2-[VAR5]
 -3-[VAR6]

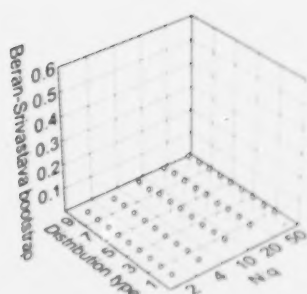
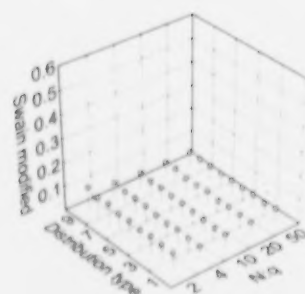
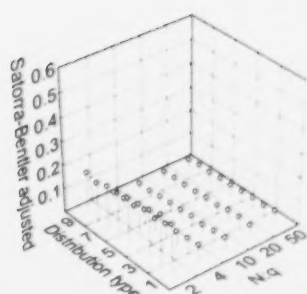
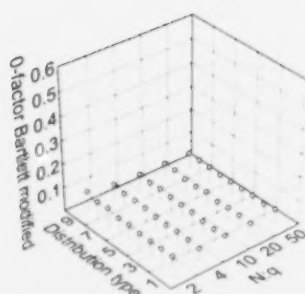
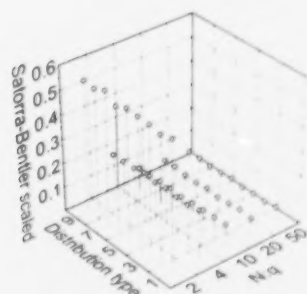
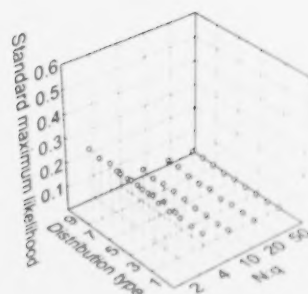
[VAR4]-1-[VAR5]
 -2-[VAR6]

[VAR5]-1-[VAR6]

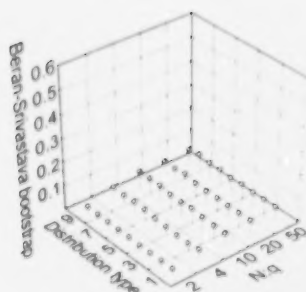
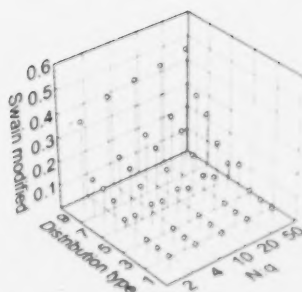
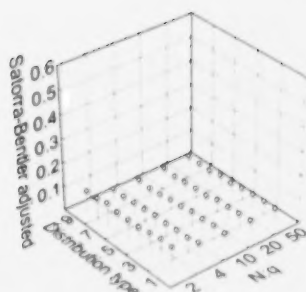
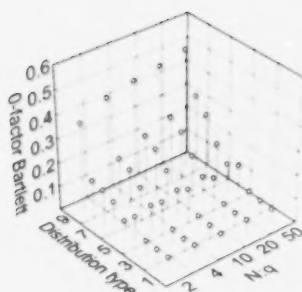
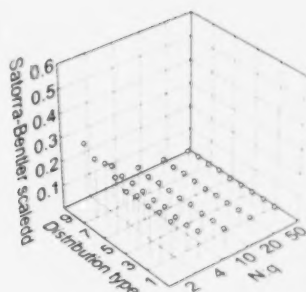
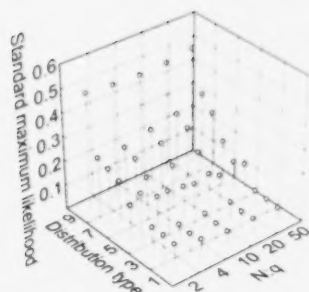
[VAR1]- 6 -[VAR1]
 [VAR2]- 7 -[VAR2]
 [VAR3]- 8 -[VAR3]
 [VAR4]- 9 -[VAR4]
 [VAR5]- 10 -[VAR5]
 [VAR6]- 11 -[VAR6]

Fig 2

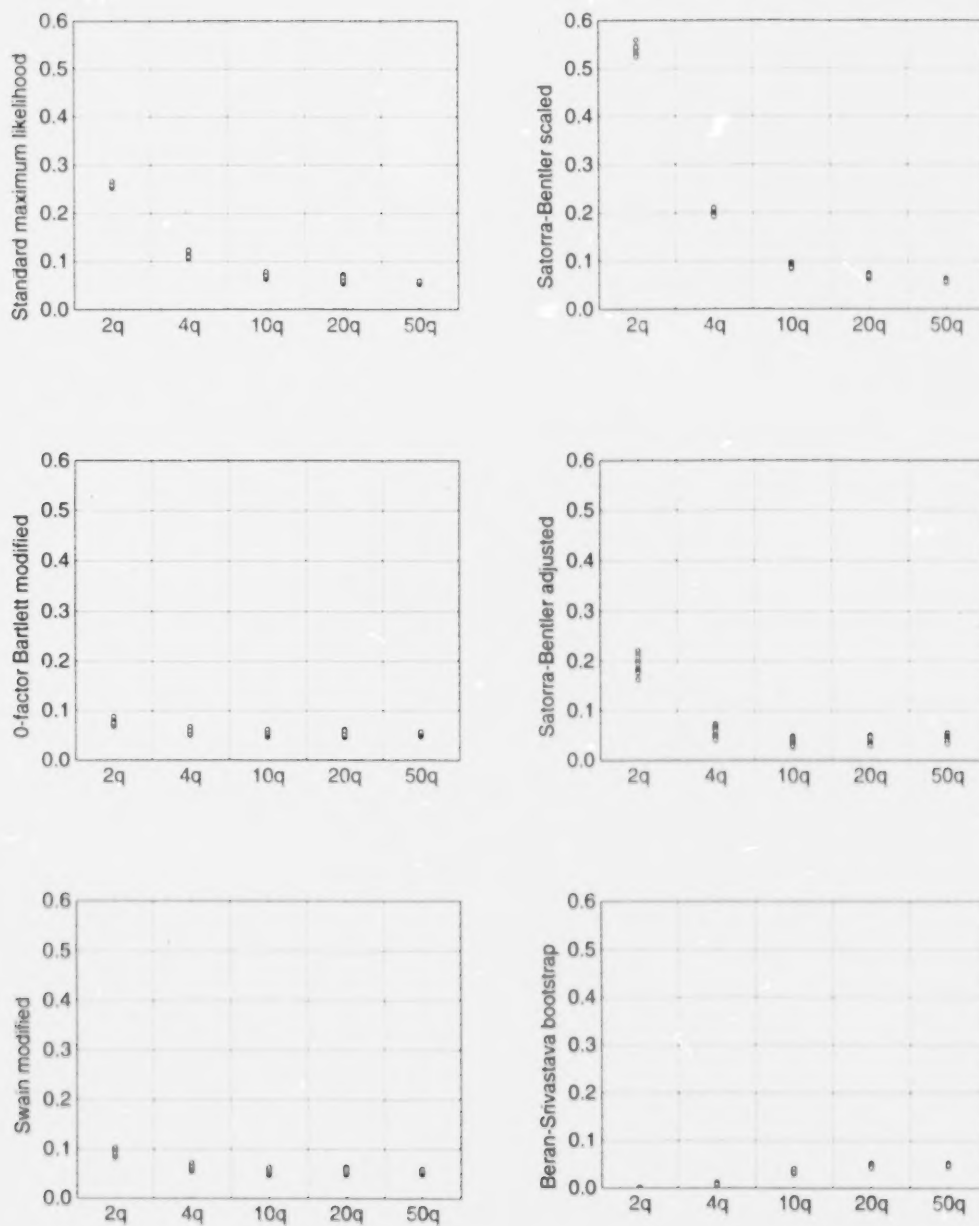
Empirical rejection rates as a function of distribution type
and sample size under conditions of the uncorrelated
variables model and nominal $\alpha=.05$



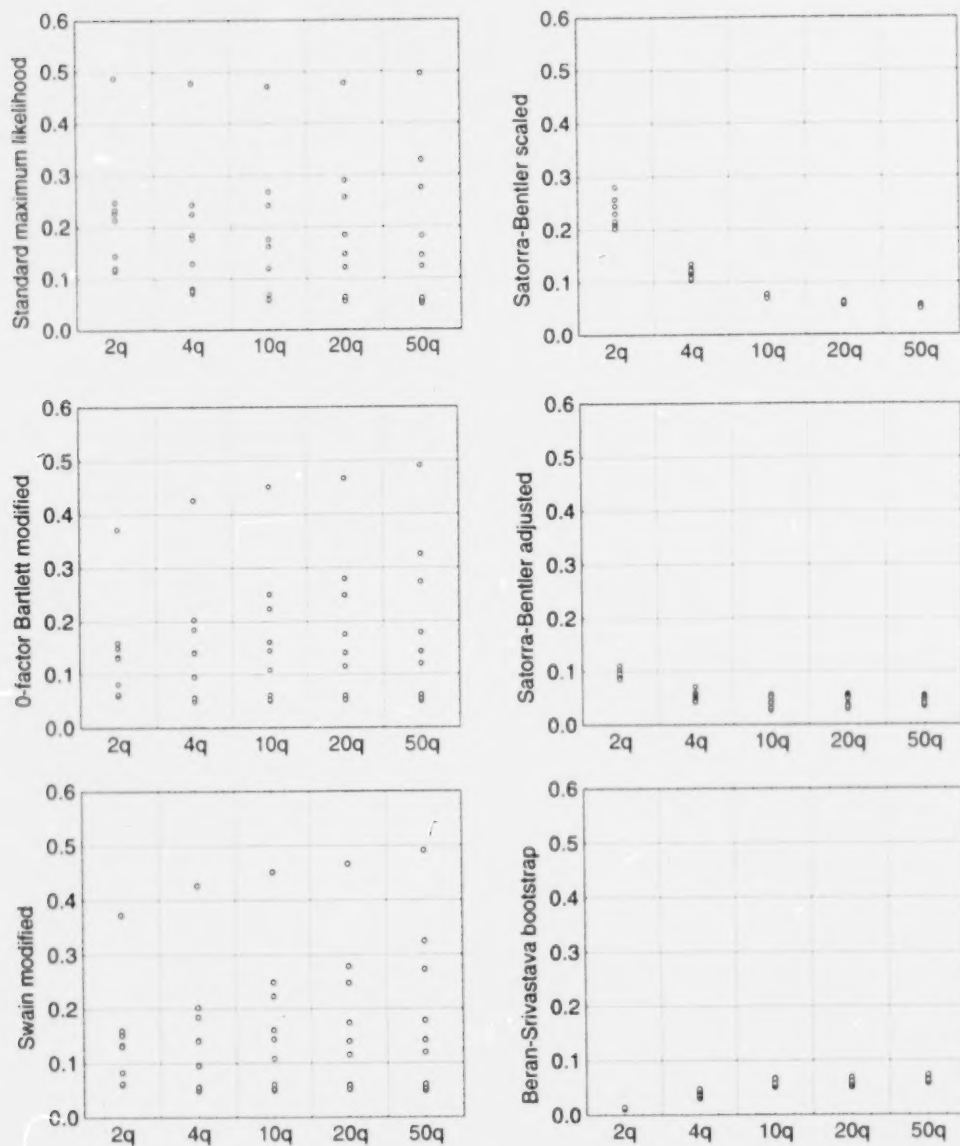
Empirical rejection rates as a function of distribution type
and sample size under conditions of the correlated
variables model and nominal $\alpha=.05$



Empirical rejection rates as a function of sample size
under conditions of the uncorrelated variables model
and nominal $\alpha=.05$

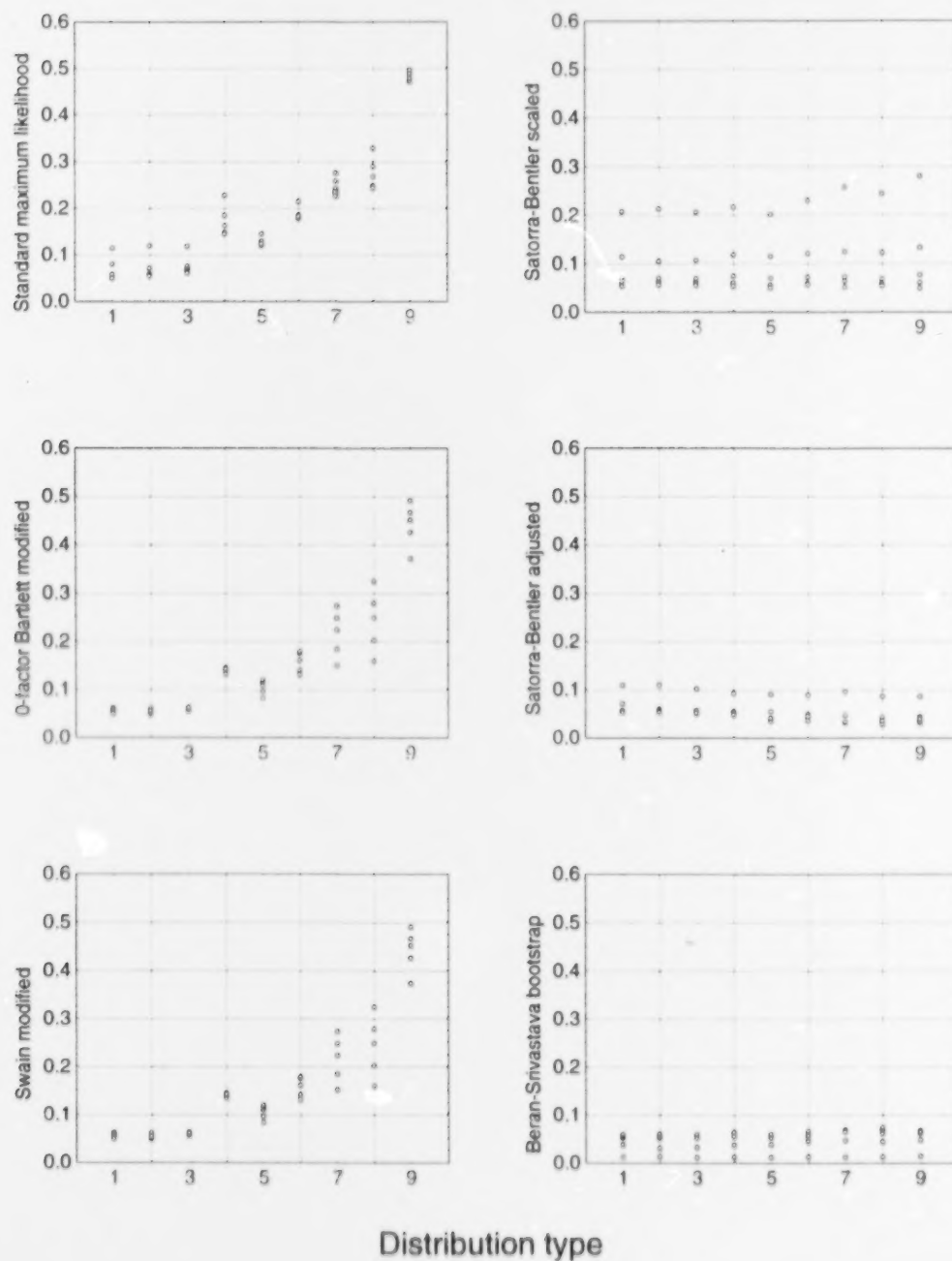


Empirical rejection rates as a function of sample size
under conditions of the correlated variables model
and nominal $\alpha = .05$

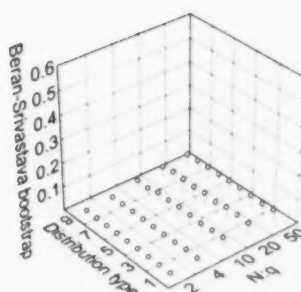
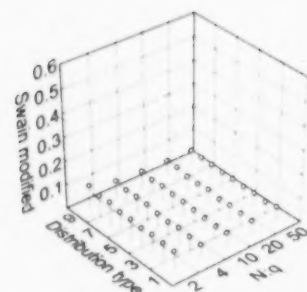
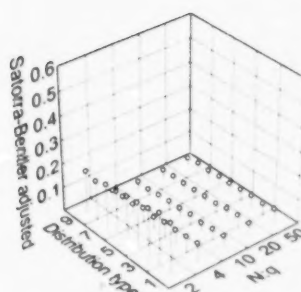
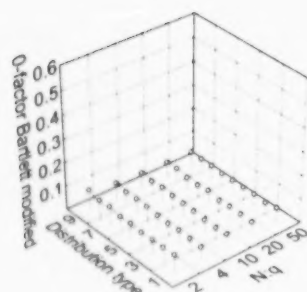
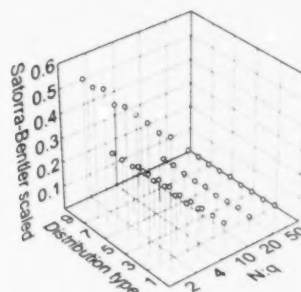
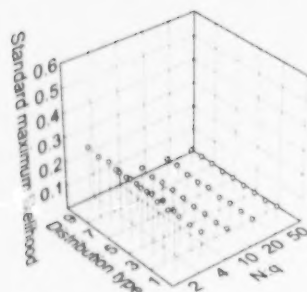


N

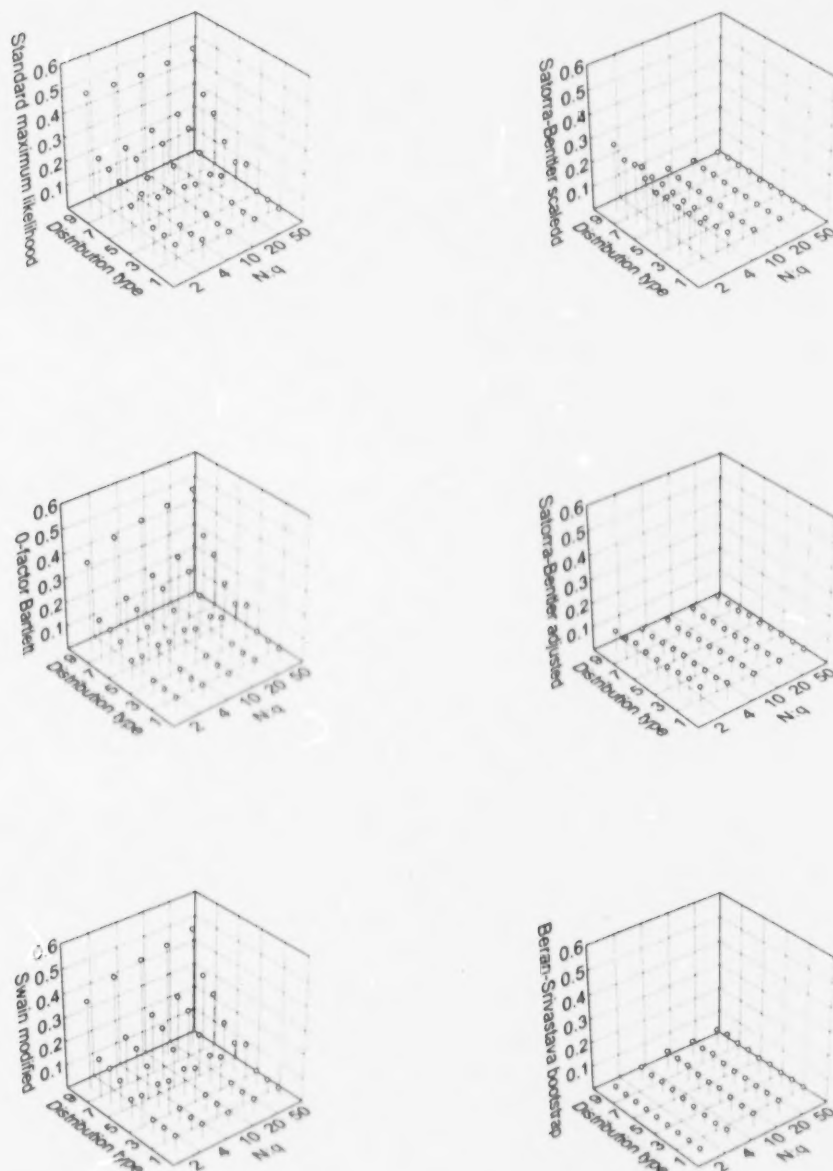
Empirical rejection rates as a function of distribution type
under conditions of the correlated variables model
and nominal $\alpha=.05$



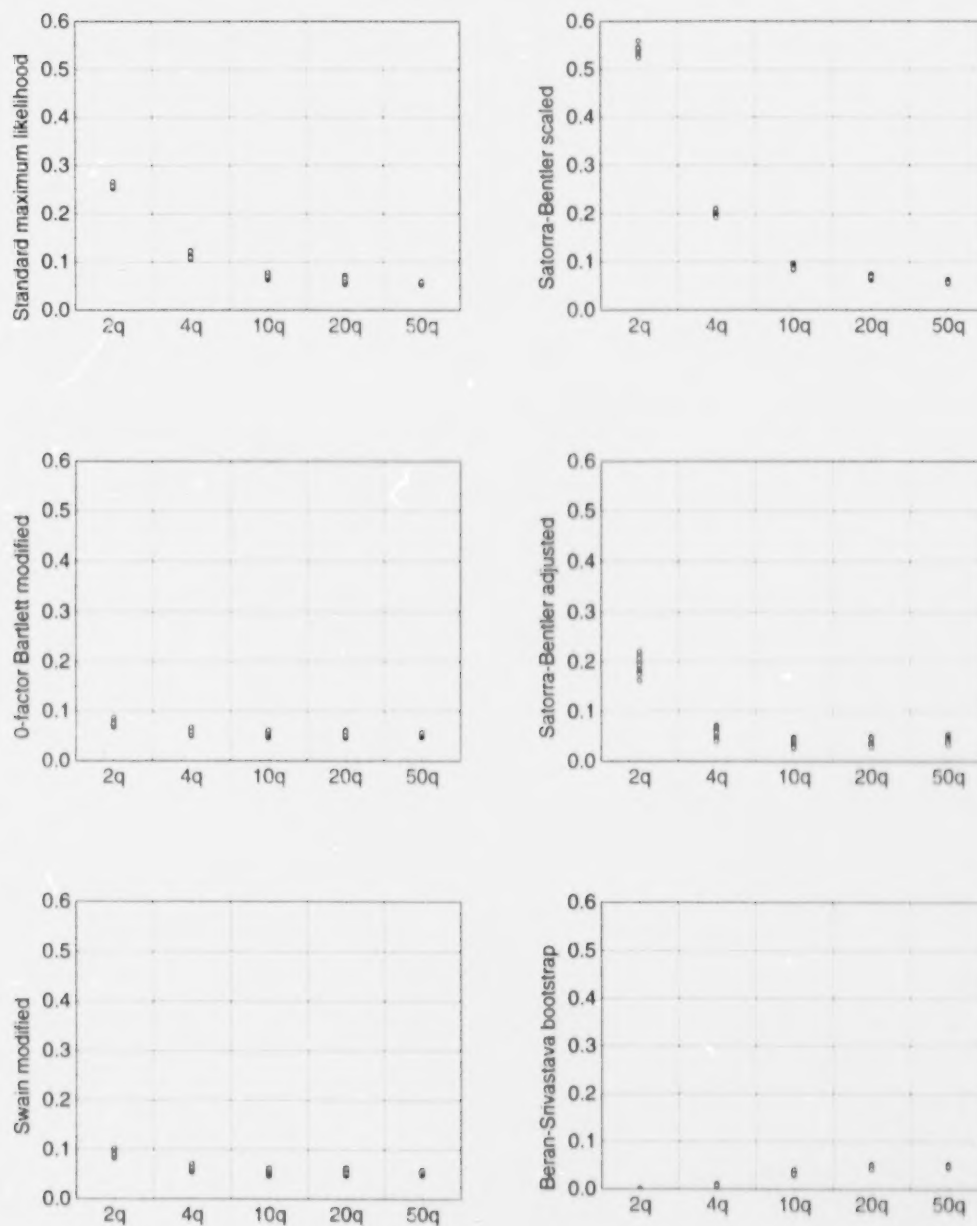
Empirical rejection rates as a function of distribution type
and sample size under conditions of the uncorrelated
variables model and nominal $\alpha=.05$



Empirical rejection rates as a function of distribution type
and sample size under conditions of the correlated
variables model and nominal $\alpha=.05$

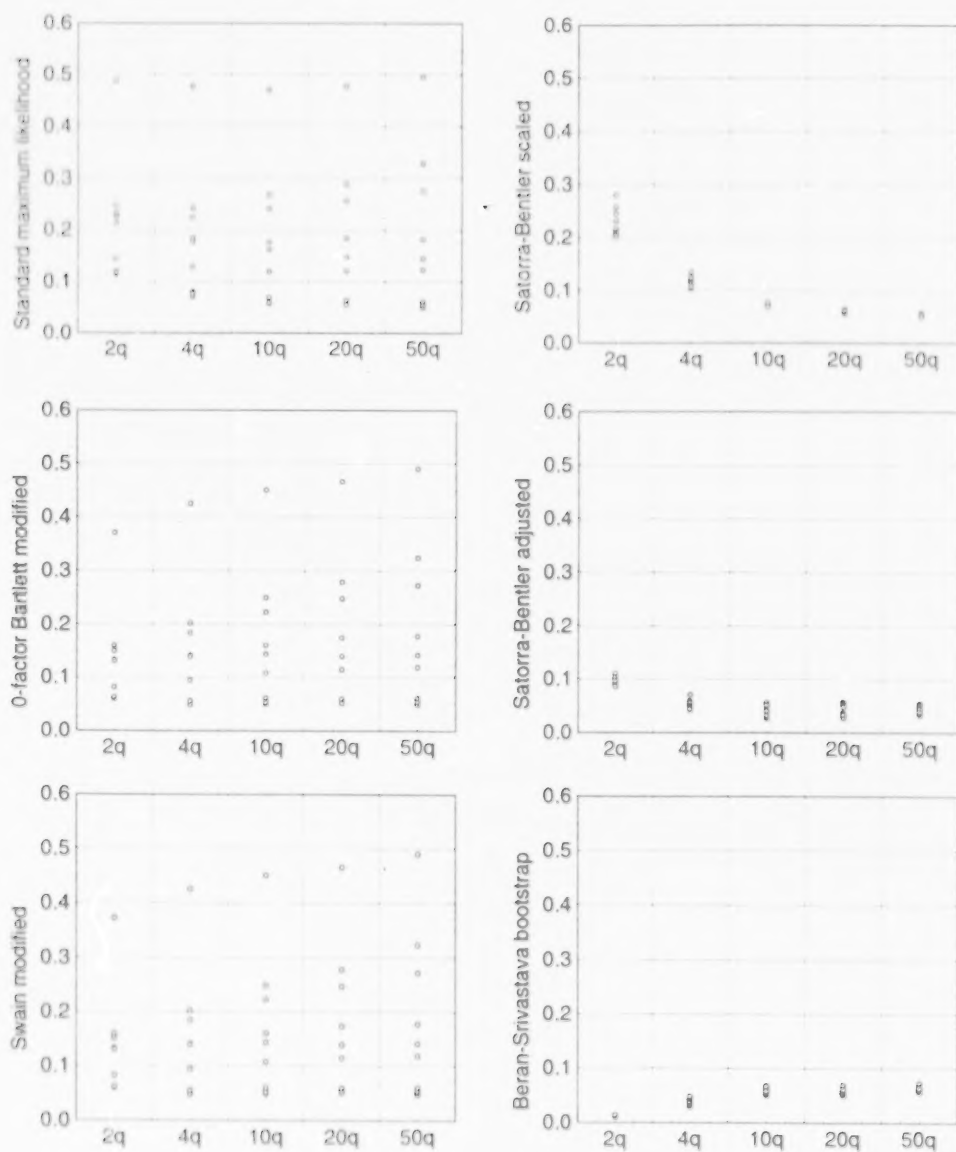


Empirical rejection rates as a function of sample size
under conditions of the uncorrelated variables model
and nominal $\alpha=.05$



N

Empirical rejection rates as a function of sample size
under conditions of the correlated variables model
and nominal $\alpha=.05$



N

Empirical rejection rates as a function of distribution type
under conditions of the correlated variables model
and nominal $\alpha=.05$

